Request to Archive
With The National Centers for Environmental Information
For National Oceanic and Atmospheric Administration (NOAA)-Cooperative Institute for Research in Environmental Sciences (CIRES) Twentieth Century Reanalysis (V2c) Provided by NOAA/OAR/ESRL

2016-06-28

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Cathy Smith

ESRL

303 497 6263

cathy.smith@noaa.gov

2. Name the organization or group responsible for creating the dataset.

NOAA/ESRL/PSD

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

The data are model outputs from an NCEP climate model run with SST and sea-ice boundary conditions. Atmospheric pressure measurements are used as inputs to the model. The model is a 3D atmospheric model with an output horizontal resolution of 2x2 degrees and 24 pressure levels. The model produces 56 ensemble members at each time step. There are multiple pressure levels variables including wind and temperature. There are also many single level variables including precipitation, temperature, radiative fluxes, winds, clouds, among many others. Data is available 4-8times a day and daily and monthly values are also created. Climatologies on all 3 times scale are available. In addition to the ensemble mean, there are ensemble spreads for sub-daily and an average over months.

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)

From 1851-01-01 to 2014-12-31

5. Edition or version number(s) of the dataset:

V2c

6. Approximate date when the dataset was or will be released to the public:

2015-02-15

7. Who are the expected users of the archived data? How will the archived data be used?

Researchers (high school to universities) and those interested in historic data (historians, farmers, fishermen, flood/drought experts, etc.)

Also for teaching purposes in classes from high School to universities). As boundary conditions for some other models of the ocean/air/land.

8. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?

There are several peer reviewed publications on the dataset.

9. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?

The 20th century reanalysis project is an extremely long (from 1851 to near present) 3-D atmospheric dataset . For studies of climate and old weather events, it is invaluable.

10. List the input datasets and ancillary information used to produce the data.

The ISPD 3.2.9 pressure databank, the COBE-SST2 sea ice dataset, the (SODAsi) version 2 until 2013, and the NOAA OI SST after 2012.

The model run was the NCEP GFS 2008ex.

11. List web pages and other links that provide information on the data.

http://www.esrl.noaa.gov/psd/data/20thC_Rean/

http://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2c.html

http://portal.nersc.gov/project/20C_Reanalysis/ for individual ensemble members,

12. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.

1. Data files are in netCDF-4 and are CF compliant. We have ISO metadata here
    http://www.esrl.noaa.gov/psd/mddb2/WAF/NOAA-CIRES_20th_Century_Reanalysis_V2c.xml
    The model is run, output is interpolated to a 2x2 grid (or about 2x2 gaussian grid) and for pressure, 24 levels. The 56 ensemble members are averaged to produce an ensemble mean and the standard deviation of the members are stored. We make available each of these (mean and spread) for the NCEI archive.

13. Indicate the data file format(s).

1. netCDF-4

14. Are the data files compressed?

netCDF-4/HDF5 compression

15. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).

For daily and sub-daily files, the data are annual. var.(optional level).yyyy.nc

All files for monthly have all months and one variable

var.optionalelvel.mean.nc

16. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

All files are

ftp ftp.cdc.noaa.gov

cd /Datasets/20thC_ReanV2c/

we also have TDS access. We do not aggregate them in time at this point (daily/sub-daily)

http://www.esrl.noaa.gov/psd/thredds/catalog/Datasets/20thC_ReanV2c/catalog.html

17. What is the total data volume to be submitted?

Historic Data: all historic data or data submitted as a completed collection.

Total Data Volume: 4.7TB

Number of Data Files: 51337

18. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.

Possibly though not too likely. If we find errors or issues, we may fix those files. This dataset version will not be appended to in time.

19. Describe the server that will connect to the ingest server at NCEI for submitting the data.

Physical Location: Boulder, CO

System Name: ftp.cdc.noaa.gov

System Owner: DOC/NOAA/ESRL/PSD

Additional Information: /Datasets/20CRV2c

20. What are the possible methods for submitting the data to NCEI? Select all that apply.

1. FTP PULL

TDS

21. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.

1. User interface to order and stage data for download

2. Direct download links

3. Advanced web services (e.g., THREDDS Catalog Service)

22. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?

Constraint Type	Description
Use	Please cite publication.

23. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.

the data is used extensively in research and other projects. If PSD loses the ability to distribute the data, it is still very important.

http://www.esrl.noaa.gov/psd/data/20thC_Rean/pubs/

http://www.esrl.noaa.gov/psd/data/20thC_Rean/news/

24. Are the data archived at another facility or are there plans to do so? Please explain.

It is archived at here at NOAA/ESRL/PSD. I believe the UKMet office may make it available for researchers there.

25. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?

I asked last week. Nothing was agreed on before that. Cost was never mentioned.

26. Do you have a data management plan for your data?

At OAR

27. Have funds been allocated to archive the data at NCEI?

No

28. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.

The NOAA-CIRES Twentieth Century Reanalysis Project version 2c used resources of the National Energy Research Scientific Computing Center managed by Lawrence Berkeley National Laboratory which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

29. Is there a desired deadline for NCEI to archive and provide access to the data?

No deadlines for archive or access.

30. Add any other pertinent information for this request.

Please contact me about cost. We have no funds allocated at this time for NCEI storage.